Media exposure and use of modern family planning

Pakistan Demographic Health Survey Data

2023-10-26

Table of contents

Prof.Dr. Zahid Asghar, School of Economics, QAU

## Research Question

#### Does modern contraceptive use among women currently differ by family planning media exposure after controlling for socio-demographic variables

I have written [PDHS using R](httsp:/zahidasghar.com/pdhsurvey/ir) few weeks back to indicate that how one can easily use DHS data with R. This post is next phase of the earlier post. I am using in this post is IR which is the individual (women’s) recode file. For more details about files you can visit [dhs](dhsprogram.com). USAID main sponsoring agency of DHS data has recently started using open community software for DHS data. I am using it for Pakistani data 2017-18 to explore whether media has a role in use of modern family planning methods or not.

## DHS data in R

Please have the dataset downloaded and ready to use. I shall use the IR file.You can download the model dataset [here](https://www.dhsprogram.com/data/Model-Datasets.cfm) For the code below to work, you must save the dataset in the same folder as your R scripts.

Install and load the packages you need. You only need to install packages once. After installing a package you can load the package using the library command.

### Load relevant packages

library(expss)   
  
library(writexl)  
library(tidyverse) # its an umbrella library containing 9 important and widely used packages of R  
library(naniar) # to use replace\_with\_na function  
library(haven) # to read STATA/SPSS data  
library(here) # to check what is your file path  
library(labelled) # to get labels of vairables  
library(pollster)  
library(gt) # to get tables  
library(gtExtras)  
library(survey) # For analysis of complex survey data sets  
here::here() # check your path/directory and make sure your data is in this directory

[1] "D:/RepTemplates/pdhsurvey"

## Load data

As I have my data in a folder data , so I am giving a parth inside here of the relevant folder and then data.

# open your dataset  
#pkir <- read\_dta(here("data","PKIR71FL.dta"))  
#   
# save(pkir, file = 'data/pkir.RData') #you can save data in R. It will reduce file size and if its correctly save, you can just click on data and it will be uploaded with the command as follows:  
  
load("D:/RepTemplates/pdhs/data/pkir.RData")  
  
pkir |> dim()

[1] 15068 5331

## Modern contraceptive use

To recode modern contraceptive use into a binary variable, first I check category labels of v313 representing methods used for family planning.

print\_labels(pkir$v313) # print\_labels command for variable labels

Labels:  
 value label  
 0 no method  
 1 folkloric method  
 2 traditional method  
 3 modern method

## Outcome variable

Recode v313 as 1 if female is using modern method and 0 if otherwise. For this I use mutate for creating a new variable modfp. To create survey weight variables wtand finally a table of summary statistics is calculated.

# creating a new categorical variable using mutate and case\_when commands  
  
pkir <- pkir |>  
 mutate(modfp =  
 case\_when(v313 == 3 ~ 1,  
 v313 < 3 ~ 0 )) |>  
 set\_value\_labels(modfp = c(yes = 1, no = 0)) |>  
 set\_variable\_labels(modfp ="Currently used any modern method")  
  
# creating the sampling weight variable.   
pkir <- pkir |> mutate(wt=v005/1000000)  
#pkir$wt <- pkir$v005/1000000  
  
# check with final report Table 7.3  
topline(pkir, modfp, wt ) |> gt() |> gt\_theme\_538()

| Currently used any modern method | Frequency | Percent | Valid Percent | Cumulative Percent |
| --- | --- | --- | --- | --- |
| no | 9353.449 | 75.65067 | 75.65067 | 75.65067 |
| yes | 3010.551 | 24.34933 | 24.34933 | 100.00000 |

In the above code we use NA\_integer\_ to replace as missing any values where women are not currently in use. There are 24% females who are using modern family planning methods.

## Socio-demographic variables

#### age

For age use v013.

### education: recode v106 to combine secondary and higher categories

pkir <- pkir %>%  
 mutate(edu =  
 case\_when(v106 ==0 ~ 0,  
 v106 ==1 ~ 1,  
 v106 >=2 ~ 2)) %>%  
 set\_value\_labels(edu = c(none = 0, primary = 1, "sec+" = 2)) %>%  
 set\_variable\_labels(edu ="education level")

#### wealth quintile: use v190

#### place of residence: use v025

#### region: use v024

#### Family planning (FP) messages exposure

#### Heard a family planning message from radio, TV or paper

#### Did not hear a family planning message from radio, TV or paper

pkir <- pkir %>%  
 mutate(fpmessage =  
 ifelse(v384a==1 | v384b==1 | v384c==1, 1, 0)) %>%  
 set\_value\_labels(fpmessage = c(yes = 1, no = 0)) %>%  
 set\_variable\_labels(fpmessage = "Exposed to TV, radio, or paper media sources")

## Dealing with missing values

For all variables, replace as missing if woman is not currently in a union, and reduce your dataset to mydata

pkir <- pkir |> mutate(wt=v005/1000000)  
vars <- c("modfp","v013","edu","fpmessage","v502", "v190","v025","v024","v005", "v007", "wt","v021","v022","v001","v002","v003")  
print\_labels(pkir$v502)

Labels:  
 value label  
 0 never in union  
 1 currently in union/living with a man  
 2 formerly in union/living with a man

mydata <- pkir %>%  
 filter(v502 == 1) %>%  
 select(all\_of(vars))

## Set the survey design using svydesign

# attaching data   
attach(mydata)

Setting the survey design using the svydesign command from the survey package the survey design will be saved in an object named mysurvey, you can change this name to another name

mysurvey<-svydesign(id=mydata$v021, data=mydata, strata=mydata$v022, weight=mydata$wt, nest=T)  
options(survey.lonely.psu="adjust")

## Descriptive Statistics and Crosstabulations

### Descriptive table

This table will include all the variables in your analysis (would be your Table 1 of your results). The variables are tabulated among women currently in a union since this is our analytical sample you can use the following code for checking the proportions of a variable

prop.table(svytable(~modfp, mysurvey))

modfp  
 0 1   
0.7498149 0.2501851

To export a table of the weighted percentages of all your variables you can use the following code

# dummy var for all women currently in a union  
mydata <- mydata %>% mutate(fp\_all = case\_when(v007>0 ~ "all"))  
  
# set expss package options to show one decimal place  
expss\_digits(digits=1)  
  
tab<- mydata %>%   
 cross\_cpct(  
 cell\_vars = list(modfp,v013,edu,fpmessage,v190,v025,v024),  
 col\_vars = list(fp\_all),  
 weight = wt,  
 expss\_digits(digits=1))   
tab |> gt()

| row\_labels | all |
| --- | --- |
| 1|Currently used any modern method|no | 74.981488 |
| 1|Currently used any modern method|yes | 25.018512 |
| 1|Currently used any modern method|#Total cases | 11902.000000 |
| 1|age in 5-year groups|15-19 | 5.007945 |
| 1|age in 5-year groups|20-24 | 15.676015 |
| 1|age in 5-year groups|25-29 | 21.078768 |
| 1|age in 5-year groups|30-34 | 19.811211 |
| 1|age in 5-year groups|35-39 | 17.270140 |
| 1|age in 5-year groups|40-44 | 11.185964 |
| 1|age in 5-year groups|45-49 | 9.969957 |
| 1|age in 5-year groups|#Total cases | 11902.000000 |
| 1|education level|none | 48.791616 |
| 1|education level|primary | 16.459991 |
| 1|education level|sec+ | 34.748393 |
| 1|education level|#Total cases | 11902.000000 |
| 1|Exposed to TV, radio, or paper media sources|no | 75.917559 |
| 1|Exposed to TV, radio, or paper media sources|yes | 24.082441 |
| 1|Exposed to TV, radio, or paper media sources|#Total cases | 11901.000000 |
| 1|wealth index combined|poorest | 18.218334 |
| 1|wealth index combined|poorer | 19.420927 |
| 1|wealth index combined|middle | 20.346496 |
| 1|wealth index combined|richer | 20.917925 |
| 1|wealth index combined|richest | 21.096318 |
| 1|wealth index combined|#Total cases | 11902.000000 |
| 1|type of place of residence|urban | 36.766003 |
| 1|type of place of residence|rural | 63.233997 |
| 1|type of place of residence|#Total cases | 11902.000000 |
| 1|region|punjab | 53.055221 |
| 1|region|sindh | 23.243249 |
| 1|region|kpk | 15.598670 |
| 1|region|balochistan | 5.297360 |
| 1|region|gb | NA |
| 1|region|ict | 0.872713 |
| 1|region|ajk | NA |
| 1|region|fata | 1.932786 |
| 1|region|#Total cases | 11902.000000 |

Note that this table gives you weighted percentages but does not produce confidence intervals. Crosstabulations of each variable with the outcome variables (Table 2 of your results)

### Crosstabulation of modfp (modern FP use) and place of residence (v025)

svyby(~modfp, by = ~v025 , design=mysurvey, FUN=svymean, vartype=c("se", "ci"))

v025 modfp se ci\_l ci\_u  
1 1 0.2877225 0.009322473 0.2694508 0.3059942  
2 2 0.2283598 0.009058787 0.2106049 0.2461147

## chi-square results

svychisq(~modfp+v025, mysurvey)

Pearson's X^2: Rao & Scott adjustment  
  
data: svychisq(~modfp + v025, mysurvey)  
F = 20.618, ndf = 1, ddf = 545, p-value = 6.905e-06

### To do this for several variables at once, you can do the following

# List of variables to crosstabulate with the outcome

variables <- c("v013","edu","fpmessage", "v190","v025","v024")   
  
results <- list() # Empty list to store the crosstabulation results  
  
# Loop through the variables  
# for (var in variables) {  
# # Crosstabulation using svyby  
# crosstab <- svyby(~ modfp, by = as.formula(paste("~", var)), design = mysurvey, FUN = svymean, vartype = c("se", "ci"))  
#   
# chi\_square <- svychisq(as.formula(paste("~ modfp +", var)), design = mysurvey)  
#   
# # Store results in list  
# results[[var]] <- list(crosstab = crosstab, chi\_square = chi\_square)  
# }

### Access the crosstabulation results for each variable

for (var in variables) {  
 print(paste("Crosstabulation for", var))  
 print(results[[var]])  
}

[1] "Crosstabulation for v013"  
NULL  
[1] "Crosstabulation for edu"  
NULL  
[1] "Crosstabulation for fpmessage"  
NULL  
[1] "Crosstabulation for v190"  
NULL  
[1] "Crosstabulation for v025"  
NULL  
[1] "Crosstabulation for v024"  
NULL

## Interpretation:

The results of the crosstabulation show that all variables were significantly associated with modern contraceptive use. To produce a table of just the percentages by modfp (i.e. no C.I.s)

tab1<- mydata %>%   
 cross\_rpct(  
 cell\_vars = list( v013,edu,fpmessage,v190,v025,v024,total()),  
 col\_vars = list(modfp),  
 weight = wt,  
 expss\_digits(digits=1))   
  
tab1 |> gt()

| row\_labels | Currently used any modern method|no | Currently used any modern method|yes |
| --- | --- | --- |
| 1|age in 5-year groups|15-19 | 94.09804 | 5.901957 |
| 1|age in 5-year groups|20-24 | 86.64317 | 13.356829 |
| 1|age in 5-year groups|25-29 | 79.07189 | 20.928115 |
| 1|age in 5-year groups|30-34 | 69.86187 | 30.138130 |
| 1|age in 5-year groups|35-39 | 67.34857 | 32.651428 |
| 1|age in 5-year groups|40-44 | 63.95045 | 36.049550 |
| 1|age in 5-year groups|45-49 | 74.16666 | 25.833336 |
| 1|age in 5-year groups|#Total cases | 9112.00000 | 2790.000000 |
| 1|education level|none | 78.43349 | 21.566510 |
| 1|education level|primary | 71.76301 | 28.236992 |
| 1|education level|sec+ | 71.65896 | 28.341043 |
| 1|education level|#Total cases | 9112.00000 | 2790.000000 |
| 1|Exposed to TV, radio, or paper media sources|no | 76.35210 | 23.647899 |
| 1|Exposed to TV, radio, or paper media sources|yes | 70.65805 | 29.341950 |
| 1|Exposed to TV, radio, or paper media sources|#Total cases | 9111.00000 | 2790.000000 |
| 1|wealth index combined|poorest | 82.89327 | 17.106734 |
| 1|wealth index combined|poorer | 77.42353 | 22.576471 |
| 1|wealth index combined|middle | 73.10917 | 26.890829 |
| 1|wealth index combined|richer | 72.37341 | 27.626594 |
| 1|wealth index combined|richest | 70.29274 | 29.707262 |
| 1|wealth index combined|#Total cases | 9112.00000 | 2790.000000 |
| 1|type of place of residence|urban | 71.22775 | 28.772254 |
| 1|type of place of residence|rural | 77.16402 | 22.835982 |
| 1|type of place of residence|#Total cases | 9112.00000 | 2790.000000 |
| 1|region|punjab | 72.82300 | 27.177002 |
| 1|region|sindh | 75.58088 | 24.419122 |
| 1|region|kpk | 76.83004 | 23.169960 |
| 1|region|balochistan | 85.98576 | 14.014239 |
| 1|region|gb | NA | NA |
| 1|region|ict | 65.33211 | 34.667891 |
| 1|region|ajk | NA | NA |
| 1|region|fata | 86.30192 | 13.698080 |
| 1|region|#Total cases | 9112.00000 | 2790.000000 |
| 1|#Total | 74.98149 | 25.018512 |
| 1|#Total cases | 9112.00000 | 2790.000000 |

## Logistic Regressions

We will use the svyglm function from the survey package to fit regression models. This fits generalized linear models.

# first check for correlations   
cordata <- mydata[, variables]  
# Calculate the correlation matrix  
cor\_matrix <- cor(cordata)  
print(cor\_matrix)

v013 edu fpmessage v190 v025 v024  
v013 1.000000000 -0.1014288 NA 0.0715958 -0.05514208 -0.003774969  
edu -0.101428787 1.0000000 NA 0.5745587 -0.28870752 -0.023242999  
fpmessage NA NA 1 NA NA NA  
v190 0.071595800 0.5745587 NA 1.0000000 -0.50326790 -0.110364231  
v025 -0.055142079 -0.2887075 NA -0.5032679 1.00000000 0.013657264  
v024 -0.003774969 -0.0232430 NA -0.1103642 0.01365726 1.000000000

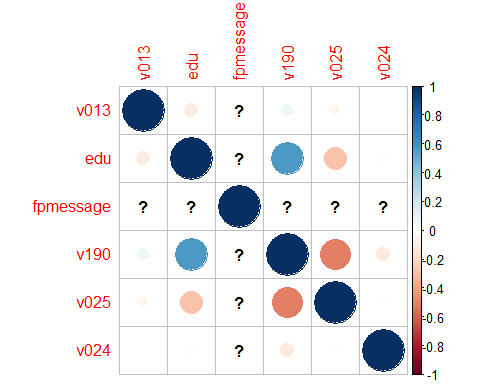
### Correlation plot

library(corrplot)

Warning: package 'corrplot' was built under R version 4.3.1

corrplot 0.92 loaded

corrplot(cor\_matrix)



There is a relatively high correlation between v025 and v190 that the researcher may need to be cautious of Unadjusted regression first we will fit an unadjusted logistic regression for family planning message exposure unadjusted with exposure to family planning messages. Ignore the warning message!

reg1 <- svyglm(modfp~ 1 + fpmessage , design=mysurvey, family=binomial(link="logit"))

Warning in eval(family$initialize): non-integer #successes in a binomial glm!

# to see the results  
library(broom)

Warning: package 'broom' was built under R version 4.3.1

tidy(reg1)

# A tibble: 2 × 5  
 term estimate std.error statistic p.value  
 <chr> <dbl> <dbl> <dbl> <dbl>  
1 (Intercept) -1.17 0.0423 -27.7 7.43e-99  
2 fpmessage 0.293 0.0659 4.45 1.09e- 5

# to get ORs  
ORreg1 <- exp(reg1$coefficients )  
ORreg1

(Intercept) fpmessage   
 0.3097217 1.3407745

## Interpretation

The unadjusted result shows that women currently in a union who have heard a FP message through one of the three media sources have almost twice the odds of using modern a contraceptive method compared to women with no FP message exposure.

#### Adjusted regression

Now we will include all the background variables with our main variable of interest note that for binary variables, you do not need to use as.factor but for categorial variables, you should add as.factor(var) to indicate that the variable is categorical and not continuous.

reg2 <- svyglm(modfp~ 1 + fpmessage + as.factor(edu) + as.factor(v013)   
 + as.factor(v190) + as.factor(v024) + as.factor(v024),   
 design=mysurvey, family=binomial(link="logit"))

Warning in eval(family$initialize): non-integer #successes in a binomial glm!

tidy(reg2)

# A tibble: 19 × 5  
 term estimate std.error statistic p.value  
 <chr> <dbl> <dbl> <dbl> <dbl>  
 1 (Intercept) -3.14 0.240 -13.1 3.58e-33  
 2 fpmessage 0.0746 0.0673 1.11 2.68e- 1  
 3 as.factor(edu)1 0.252 0.0959 2.63 8.93e- 3  
 4 as.factor(edu)2 0.222 0.0956 2.32 2.08e- 2  
 5 as.factor(v013)2 0.829 0.212 3.91 1.06e- 4  
 6 as.factor(v013)3 1.37 0.219 6.26 9.53e-10  
 7 as.factor(v013)4 1.86 0.214 8.70 7.28e-17  
 8 as.factor(v013)5 1.99 0.217 9.19 1.72e-18  
 9 as.factor(v013)6 2.17 0.226 9.61 6.51e-20  
10 as.factor(v013)7 1.68 0.234 7.19 2.82e-12  
11 as.factor(v190)2 0.321 0.113 2.83 4.83e- 3  
12 as.factor(v190)3 0.442 0.120 3.67 2.77e- 4  
13 as.factor(v190)4 0.388 0.123 3.16 1.67e- 3  
14 as.factor(v190)5 0.375 0.138 2.71 7.02e- 3  
15 as.factor(v024)2 0.00262 0.0882 0.0297 9.76e- 1  
16 as.factor(v024)3 -0.0425 0.0925 -0.459 6.46e- 1  
17 as.factor(v024)4 -0.628 0.156 -4.03 6.51e- 5  
18 as.factor(v024)6 0.232 0.104 2.24 2.57e- 2  
19 as.factor(v024)8 -0.479 0.238 -2.01 4.45e- 2

# to see the results  
summary(reg2)

Call:  
svyglm(formula = modfp ~ 1 + fpmessage + as.factor(edu) + as.factor(v013) +   
 as.factor(v190) + as.factor(v024) + as.factor(v024), design = mysurvey,   
 family = binomial(link = "logit"))  
  
Survey design:  
svydesign(id = mydata$v021, data = mydata, strata = mydata$v022,   
 weight = mydata$wt, nest = T)  
  
Coefficients: (2 not defined because of singularities)  
 Estimate Std. Error t value Pr(>|t|)   
(Intercept) -3.140442 0.239877 -13.092 < 2e-16 \*\*\*  
fpmessage 0.074562 0.067266 1.108 0.268282   
as.factor(edu)1 0.251789 0.095859 2.627 0.008932 \*\*   
as.factor(edu)2 0.221787 0.095624 2.319 0.020845 \*   
as.factor(v013)2 0.829073 0.211833 3.914 0.000106 \*\*\*  
as.factor(v013)3 1.367586 0.218567 6.257 9.53e-10 \*\*\*  
as.factor(v013)4 1.859005 0.213710 8.699 < 2e-16 \*\*\*  
as.factor(v013)5 1.994699 0.217048 9.190 < 2e-16 \*\*\*  
as.factor(v013)6 2.167888 0.225691 9.606 < 2e-16 \*\*\*  
as.factor(v013)7 1.683638 0.234007 7.195 2.82e-12 \*\*\*  
as.factor(v190)2 0.321270 0.113399 2.833 0.004828 \*\*   
as.factor(v190)3 0.441658 0.120450 3.667 0.000277 \*\*\*  
as.factor(v190)4 0.388075 0.122680 3.163 0.001671 \*\*   
as.factor(v190)5 0.375132 0.138472 2.709 0.007018 \*\*   
as.factor(v024)2 0.002617 0.088239 0.030 0.976356   
as.factor(v024)3 -0.042455 0.092459 -0.459 0.646341   
as.factor(v024)4 -0.627757 0.155646 -4.033 6.51e-05 \*\*\*  
as.factor(v024)6 0.232418 0.103851 2.238 0.025735 \*   
as.factor(v024)8 -0.478759 0.237602 -2.015 0.044533 \*   
---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
  
(Dispersion parameter for binomial family taken to be 0.9962998)  
  
Number of Fisher Scoring iterations: 5

# to get ORs  
ORreg2 <- exp(reg2$coefficients )  
ORreg2

(Intercept) fpmessage as.factor(edu)1 as.factor(edu)2   
 0.04326366 1.07741192 1.28632422 1.24830564   
as.factor(v013)2 as.factor(v013)3 as.factor(v013)4 as.factor(v013)5   
 2.29119363 3.92586312 6.41734969 7.34999249   
as.factor(v013)6 as.factor(v013)7 as.factor(v190)2 as.factor(v190)3   
 8.73980212 5.38511243 1.37887725 1.55528361   
as.factor(v190)4 as.factor(v190)5 as.factor(v024)2 as.factor(v024)3   
 1.47414056 1.45518329 1.00262018 0.95843354   
as.factor(v024)4 as.factor(v024)5 as.factor(v024)6 as.factor(v024)7   
 0.53378779 NA 1.26164758 NA   
as.factor(v024)8   
 0.61955164

## Interpretation:

After controlling for other variables, exposure to FP messages was found to be significantly associated with modern contraceptive use among women currently in a union. Women with FP message exposure had 1.5 times higher odds compared to women with no exposure in using modern contraceptive methods.